

SEQUENCE LISTING

<110> Lassen, Soren Flensted
<120> Improved proteases and methods for producing them
<130> 10495.204-US
<160> 53
<170> PatentIn version 3.3
<210> 1
<211> 1062
<212> DNA
<213> Nocardiopsis sp. NRRL 18262

<220>
<221> misc_feature
<222> (1)..(495)
<223> Encodes the pro-region shown in positions -165 to -1 of SEQ ID NO:43.

<220>
<221> misc_feature
<222> (496)..(1059)
<223> Encodes the mature region shown in positions 1-188 of SEQ ID NO:43.

<400> 1
gctactggag cattaccta gtctcctaca cctgaagcag atgcagtatc gatgcaagaa 60
gcattacaac gtgatcttga tcttacatca gctgaagctg aggaattact tgctgcacaa 120
gatacagcct ttgaagttga tgaagctgcc gctgaagcag ctggtgatgc atatgggt 180
tcagtattcg atactgaatc actcgaacct actgtacttag tgaccgatgc agcagctgtt 240
gaagctgttg aagccacagg tgcaggtaca gagctcgat cttatggtat tggatggatta 300
gatgagatcg tacaagagct taatgcagct gatgccgttc caggtgttgt tggatggtat 360
cctgatgttag caggtgatac tgggtctta gaagttcttg aaggctctgg agctgatgtt 420
tctggacttt tagcagacgc aggagtcat gcatccgcgg ttgaagtgac cacgtcagat 480
cagcctgaac tctatgccga tatcattgga ggcctagcgt acacaatggg tggtcgctgc 540
agcgttaggat ttgcagccac aaatgcagct ggacaacctg gcttcgtgac agctggacat 600
tgccggccgcg tcggtacaca ggttactatc ggcaatggaa gaggtgtctt tgagcaaagc 660
gtatattcccg ggaatgatgc tgccttcgtt agaggtacgt ccaactttac gcttactaac 720
ttagtatcta gatacaaacac tggcggatat gcaactgttag caggtcacaa tcaagcacct 780

attggctcta	gcgtctgccg	ctcagggtcg	actacaggat	ggcattgtgg	aaccattcaa	840
gctagaggtc	agagcgtgag	ctatcctgaa	ggtaccgtaa	cgaacatgac	tcgtacgact	900
gtatgtgcag	aaccaggtga	ctctggaggt	tcatatatca	gcggtacgca	agcgcaaggc	960
gttacctcag	gtggatccgg	taactgttagg	acaggtggca	caacgttcta	ccaggaagtg	1020
acaccgatgg	tgaactcttg	gggagttaga	ctccgtacat	aa		1062

<210> 2
<211> 1143
<212> DNA
<213> Artificial sequence

<220>
<223> A synthetic 10R gene (10Rsyt-15) encoding a S2A protease denoted "10R" fused by PCR in frame to the signal peptide encoding sequence of a heterologous protease, Savinase.

<400> 2	atgaagaaac	cgttgggaa	aattgtcgca	agcacccgac	tactcatttc	tgttgcttt	60
	agttcatcga	tcgcacatcgcc	tgctactgga	gcattacctc	agtctcctac	acctgaagca	120
	gatgcagtat	cgatgcaga	agcattacaa	cgtgatctt	atcttacatc	agctgaagct	180
	gaggaattac	ttgctgcaca	agatacagcc	tttgaagttt	atgaagctgc	cgctgaagca	240
	gctggtgatg	catatggtg	ttcagtattt	gatactgaat	cactcgaact	tactgtacta	300
	gtgaccgatg	cagcagctgt	tgaagcttt	gaagccacag	gtgcaggtac	agagctcgta	360
	tcttatggta	ttgatggatt	agatgagatc	gtacaagagc	ttaatgcagc	tgtatccgtt	420
	ccaggtgtag	ttggatggta	tcctgatgta	gcaggtgata	ctgttgtt	agaagtttt	480
	gaaggctctg	gagctgatgt	ttctggactt	ttagcagacg	caggagtcga	tgcattcccg	540
	gttgaagtga	ccacgtcaga	tcagcctgaa	ctctatgccc	atatcattgg	aggcctagcg	600
	tacacaatgg	gtggtcgtg	cagcgttagga	tttgcagcca	caaatgcagc	tggacaacct	660
	ggcttcgtga	cagctggaca	ttgcggccgc	gtcggcacac	aggttactat	cggcaatgg	720
	agaggtgtct	ttgagcaaag	cgtatccc	ggaaatgatg	ctgccttcgt	tagaggtacg	780
	tccaacttta	cgcttactaa	cttagtatct	agataacaaca	ctggcggata	tgcaactgta	840
	gcaggtcaca	atcaagcacc	tattggctct	agcgctgccc	gctcagggtc	gactacagga	900
	tggcattgtg	gaaccattca	agcttagaggt	cagagcgtga	gctatcctga	aggtaccgta	960
	acgaacatga	ctcgatcgac	tgtatgtcga	gaaccaggtg	actctggagg	ttcatatatc	1020
	agcggtacgc	aagcgcaagg	cgttacacta	ggtggatccg	gtaactgtag	gacaggtggc	1080

acaacgttct accaggaagt gacaccgatg gtgaactctt ggggagtttag actccgtaca 1140
taa 1143

<210> 3
<211> 8
<212> PRT
<213> Artificial sequence

<220>
<223> C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 3

Gln Ser His Val Gln Ser Ala Pro
1 5

<210> 4
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Polynucleotide encoding a C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 4
caatcgcatg ttcaatccgc tcca 24

<210> 5
<211> 4
<212> PRT
<213> Artificial sequence

<220>
<223> C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 5

Gln Ser Ala Pro
1

<210> 6
<211> 12
<212> DNA
<213> Artificial sequence

<220>
<223> Polynucleotide encoding a C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 6
caatcggtc ct 12

<210> 7
<211> 2
<212> PRT
<213> Artificial sequence

<220>
<223> C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 7

Gln Pro
1

<210> 8
<211> 6
<212> DNA
<213> Artificial sequence

<220>
<223> Polynucleotide encoding a C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 8
caaccca 6

<210> 9
<211> 1
<212> PRT
<213> Artificial sequence

<220>
<223> C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 9

Pro
1

<210> 10
<211> 3
<212> DNA
<213> Artificial sequence

<220>
<223> Polynucleotide encoding a C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 10		
cca		3
<210> 11		
<211> 45		
<212> DNA		
<213> Artificial sequence		
<220>		
<223> Primer #252639		
<400> 11		
catgtgcatg tgggtaccgc aacgttcgca gatgctgctg aagag		45
<210> 12		
<211> 44		
<212> DNA		
<213> Artificial sequence		
<220>		
<223> Primer #251992		
<400> 12		
catgtgcatg tggtcgaccg attatggagc ggattgaaca tgcg		44
<210> 13		
<211> 44		
<212> DNA		
<213> Artificial sequence		
<220>		
<223> Primer #179541		
<400> 13		
gcgttgagac gcgcggccgc gagcgccgtt tggctgaatg atac		44
<210> 14		
<211> 43		
<212> DNA		
<213> Artificial sequence		
<220>		
<223> Primer #179542		
<400> 14		
gcgttgagac agctcgagca gggaaaaatg gaaccgcttt ttc		43
<210> 15		
<211> 64		
<212> DNA		
<213> Artificial sequence		

```

<220>
<223> Primer #179539

<400> 15
ccatttgcatac agaattcaact ggccgtcggt ttacaaccat tgcgaaaaat agtcataggc      60
atcc                                         64

<210> 16
<211> 60
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #179540

<400> 16
ggatccagat ctggtagcccg ggtctagagt cgacgcggcg gttcgctgcc ggacagcaca      60

<210> 17
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #179154

<400> 17
gttgtaaaac gacggccagt gaattctgat caaatgg                                37

<210> 18
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #179153

<400> 18
ccgcgtcgac actagacacg ggtacctgat ctagatc                                37

<210> 19
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #317

<400> 19
tggcgcaatc ggtaccatgg gg                                         22

```

```

<210> 20
<211> 40
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #139 NotI

<400> 20
catgtgcatt cggccgcatt aacgcgttgc cgcttctg 40

<210> 21
<211> 7443
<212> DNA
<213> Artificial sequence

<220>
<223> Sequence of plasmid pMB1508

<400> 21
tcgcgcgtt cggatgac ggtaaaacc tctgacacat gcagctccc gagacggta 60
cagcttgtct gtaagcgat gccggagca gacaagcccg tcagggcgcg tcagcgggt 120
ttggcgggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatg 240
cg gtgtgaaata ccgcacagat gcgttaaggag aaaataccgc atcaggcgcc
atcgccatt caggctgcgc aactgttggg aaggcgatc ggtgcgggcc tcttcgctat 300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttggta acgccagggt 360
tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cgataaaagt gcttttttg 420
ttgcaattga agaattatta atgttaagct taattaaaga taatatctt gaattgtaac 480
gccctcaaa agtaagaact acaaaaaaag aatacg 540
ttttagat atagaaat atcgatccat atgatccat atgatccat atgatccat
tcttcagatt acaaataat tcggacggac tctaccta atgcttatct aactatagaa 600
tgacatacaa gcacaac 660
ctt gtttgcgtt gaaatattga aaatataact accaatgaac ttgttcatgt
gaattatcgc tgtat 720
tttctcaatt caatataaa tatgccaata cattgttaca
agt 780
agaaatt aagacaccct tgatagcctt actataccta acatgatgta gtat 840
aaatgtaaa tatatttata 840
ataagaagcg acttatttata aatcattaca tattttctat
ttggaatgat taagattcca atagaatagt gtataaatta tttatcttga aaggaggat 900
gcctaaaaac gaagaacatt aaaaacat 960
at atttgcaccg tctaattggat ttatgaaaa
tcattttatc agttgaaaa ttatgttata tggagctctg aaaaaaagga gaggataaag 1020
aatgaagaaa ccgttggga aaattgtcgc aagcaccgca ctactcat 1080
tttgcctt

```

tagtcatcg atcgcatcg	ctgctgaaga agcaaaagaa	aaatatttaa ttggctttaa	1140
tgagcaggaa gctgtcagt	g agtttgta acaagtagag	gcaaatgacg aggtcgccat	1200
tctctctgag	gaagaggaag tcgaaattga	attgcttcat gaatttggaaa cgattcctgt	1260
tttatccgtt	gagttAACCC	cagaagatgt ggacgcgc tt gaactcgatc cagcgatttc	1320
ttatattgaa	gaggatgcag aagtaacgac	aatggcgaa tcggtaccat ggggtatatc	1380
aacgcgttaa	tccgcggata tatagcggcc	gcagatctgg gaccaataat aatgactaga	1440
gaagaaagaa	tgaagattgt tcatgaaatt	aaggaacgaa tattggataa agtggat	1500
ttttaaaata	tatatttatg ttacagtaat	attgactttt aaaaaaggat tgattcta	1560
gaagaaagca	gacaagtaag ctcctaaat	tcactttaga taaaaattta ggaggcatat	1620
caaataact	ttaataaaat tgatTTAGAC	aattggaaga gaaaagagat atttaatcat	1680
tatttgaacc	aacaaacgac tttagtata	accacagaaa ttgatattag tgTTTatac	1740
cgaacataa	aacaagaagg atataaattt	taccctgc ttatTTCTT agtgacaagg	1800
gtgataaact	caaatacagc tttagaact	ggttacaata gcgacggaga gttaggttat	1860
tggataagt	tagagccact ttatACAATT	tttgatggtg tatctaaaac attctctgg	1920
atttggactc	ctgtAAAGAA tgacttcaaa	gagttttatg atttatacct ttctgtatgt	1980
gagaaatata	atggTCGGG gaaattgtt	cccAAACAC ctatacctga aaatgcttt	2040
tctcttcta	ttattccatg gacttcattt	actgggtta acttaaatat caataataat	2100
agtaattacc	ttctacccat tattacagca	ggAAAATTCA ttaataaagg taattcaata	2160
tatTTACCGC	tatcttaca ggtacatcat	tctgtttgtg atggTTATCA tgcaggattt	2220
tttatgaact	ctattcagga attgtcagat	aggcctaattg actggcttt ataataatgag	2280
ataatGCCGA	ctgtacttt tacagtcgg	tttctaacga tacattaata ggtacgaaaa	2340
agcaactttt	tttgcgtta aaaccagtca	taccaataac ttaagggtaa ctgcctcgc	2400
cggaaagagc	gaaaatgcct cacatttgc	ccacctaaaa aggagcgatt tacatatgag	2460
ttatgcagtt	tgtagaatgc aaaaagtga	atcagctgga ctAAAAGGGG ccgcagagta	2520
gaatggaaaa	ggggatcgga aaacaagtat	ataggaggag acctatttgg ggcTTcagaa	2580
aaagacgcag	gaaaacagtc agcagtaaa	cttgcatttgc tgcttattac tgtcgctgt	2640
ggactaatca	tctggTTTAT	tccgcgtccg tccggacttg aacctaaagc ttggcatttgc	2700
tttgcgattt	ttgtcgcaac aattatcg	tttgcatttgc aacccTTGCC aatgggtgca	2760
attgcaattt	ttgcatttgc ggttactgca	ctaactggaa cactatcaat tgaggataca	2820

ttaagcggat tcgggataa gaccatttg cttatcgta tcgcattctt tatttcccg	2880
ggatttatca aaaccggctc cggtgcgaga atttcgtatg tattcggtca gaaattcgga	2940
aaaaaaaaacc ttggactttc ttattcactg ctattcagtg attaataact ttcacctgct	3000
atccaagta atacggcgcg tgcaggaggc attatatttc ctattatcag atcattatcc	3060
gaaacattcg gatcaagccc ggcaaattgga acagagagaa aaatcggtgc attcttatta	3120
aaaaccgggtt ttcagggaa tctgatcaca tctgctatgt tcctgacagc gatggcggcg	3180
aacccgctga ttgccaagct gcccatgat gtcgcagggg tggacttaac atggacaagc	3240
tggcaattt ccgcgattgt accgggactt gtaagcttaa tcatcacgcc gcttgtgatt	3300
tacaaactgt atccggcga aatcaaagaa acacccgatg cggcggaaat cgcaacagaa	3360
aaactgaaag aaatggacc gttcaaaaaa tcggagctt ccatggttat cgtgttctt	3420
ttggtgcttg tgctgtggat tttggcggc agcttcaaca tcgacgctac cacaaccgca	3480
ttgatcggtt tggccgttct cttattatca caagttctga cttggatga tatcaagaaa	3540
gaacagggcg cttggatac gtcacttgg tttggcggc ttgtcatgct cgccaacttc	3600
ttgaatgaat taggcatggt gtctgggttc agtaatgcc a taaaatcatc cgtatcaggg	3660
ttctcttggcatt catcattttt attgttgtgt attattactc tcactatttc	3720
tttgcaggatg cgacagccca catcagtgcg atgtattcag cattttggc tgtcgctg	3780
gcagcggcg caccggcgct ttagcagcg ctgaggctcg cggtcatcag caacctgttc	3840
gggtcaacga ctcactacgg ttctggagcg gctccggctc tttcgagc aggctacatc	3900
ccgcaaggca aatggggcattt atcctgtcga ttgttcatat catcgatgg	3960
cttgcgtatcg gcggattatg gtggaaatgt ctaggaatat ggttagaaaga aaaaggcaga	4020
cgcggctgc cttttttt tttcactcct tcgtaaagaaa atggattttt aaaaatgaga	4080
aaattccctg tgaaaaatgg tatgtatctag gttagaaagga cggctggcgtc tgggtgaaa	4140
aagcggttcc attttccct gcaaacaaaa ataatggggc tgattggc tctgctggc	4200
tttgcattt gttgtgtac cattacgtta gccgttcagc atacacaggg agaacggaga	4260
caggcagagc agctggcggt tcaaacggcg agaaccattt cctatatgcc gcccgttaaa	4320
gagctcattt agagaaaaga cggacatgcg gtcagacgc aagaggtcat tgaacaaatg	4380
aaagaacaga ctgggtcggt tgccattttt gtttgaacg aaaaaggaga cattcgacg	4440
gcctctggaa aaagcggatt aaagaaaactg gagcgcagca gagaaatttt gtttggcggt	4500

tcgcatgtt	ctgaaacaaa	agcggatgga	cgaagagtga	tcagaggag	cgcgccatt	4560
ataaaagaac	agaaggata	cagccaagt	atcggcagcg	tgtctgttga	tttctgcaa	4620
acggagacag	agcaaagcat	caaaaagcat	ttgagaaatt	tgagtgtat	tgctgtgctt	4680
gtactgctgc	tcggatttat	tggcgccgcc	gtgctggcga	aaagcatcag	aaaggatacg	4740
ctcgggctt	aaccgcatt	gatcgcggct	ctatatcg	agaggaacgc	aatgc	4800
gcgattcgag	aaggattat	tgccaccaat	cgtgaaggcg	tcgtcaccat	gatgaacgta	4860
tcggcgccg	agatgtgaa	gctgcccag	cctgtatcc	atcttccat	agatgacg	4920
atgccggag	cagggctgat	gtctgtgctt	aaaaaaggag	aaatgctgc	gaaccaggaa	4980
gtaagcgtca	acgatcaagt	gtttattatc	aatacgaag	tatgatca	aggcggcag	5040
gcgtatgg	ttgtcgtag	cttcagggag	aaaacagagc	tgaagaagct	gatcgacaca	5100
ttgacagagg	ttcgcaaata	ttcagaggat	ctcagggcgc	agactcatg	atttcaat	5160
aagctttag	cgattttagg	gctgcgtcg	cctgcaggca	tgcaagctt	gcgtatcat	5220
ggtcata	gtttcctgt	tgaaattgtt	atccgctcac	aattccacac	aacatacgag	5280
ccgaaagcat	aaagttaaa	gcctgggt	cctaattg	gagctaactc	acattaatt	5340
cgttgcgtc	actgcccgt	ttccagtcg	gaaacctgtc	gtgccagct	cattaatgaa	5400
tcggccaac	cgcgggaga	ggcggttgc	gtattggcg	ctttccgct	tcctcgctca	5460
ctgactcgct	gwgctcg	gttcggctgc	ggcgagcg	atcagctcac	tcaaaggcgg	5520
taatacgg	atccacagaa	tcagggata	acgcaggaaa	gaacatgt	gcaaaaggcc	5580
agcaaaaggc	caggaaccgt	aaaaaggccg	cgttgctgg	gttttccat	aggctccg	5640
ccccgtacga	gcatcacaaa	aatcgacgt	caagtca	gtggcgaaac	ccgacaggac	5700
tataaagata	ccaggcg	ccccctggaa	gctccctcg	gctctc	gttccgaccc	5760
tgccgcttac	cgatcac	tccgc	tcccttc	taactatcg	cttgcata	5820
gctcacgt	tagtatctc	attcgg	aggcgttc	ctccaagct	ggctgtgt	5880
acgaacccc	cgttca	gaccgctgc	ccttatccg	taactatcg	cttgagtcc	5940
acccggtaag	acacgactt	tcgca	actgg	cagcagcc	tggtaacagg	6000
cgaggtatgt	aggcgg	gtct	acagat	ttct	tgaagtgg	6060
gaaggacagt	atttgg	tgc	tgat	ccagt	tac	aaaagagtt
gtagcttt	atccgg	caaaccac	ctgg	tagc	gg	tggtttttt
acagattac	g	aaaggatctc	aaga	agatcc	ttt	gatctt
g	g	aaaa	aaaggatctc	aaga	agatcc	tctacggg

ctgacgctca	gtggAACGAA	aactcacgtt	aagggatttt	ggtcatgaga	ttatcaaaaa	6300	
ggatcttcac	ctagatcctt	ttaaattaaa	aatgaagttt	taaatcaatc	taaagtatat	6360	
atagagtaaac	ttggTCTGAC	agttaccaat	gcttaatcag	tgaggcacct	atctcagcga	6420	
tctgtctatt	tcgttcatcc	atagttgcct	gactccccgt	cgtgtagata	actacgatac	6480	
gggaggggctt	accatctggc	cccagtgctg	caatgatacc	gcgagaccca	cgctcaccgg	6540	
caccggattt	atcagcaata	aaccagccag	ccggaaggc	cgagcgcaga	agtggtcctg	6600	
caactttatc	cgcctccatc	cagtctatta	attgttgccg	ggaagctaga	gtaagtagtt	6660	
cgtcagttaa	tagttgcgc	aacgttgtt	ccattgctac	aggcatcgtg	gtgtcacgct	6720	
cgtcgtttgg	tatggcttca	ttcagctccg	gttcccaacg	atcaaggcga	gttacatgat	6780	
cccccatgtt	gtgcaaaaaa	gcggtagt	ccttcggtcc	tccgatcg	tttgcagaagta	6840	
agttggccgc	agtgttatca	ctcatggta	tggcagcact	gcataattct	cttactgtca	6900	
tgccatccgt	aagatgctt	tctgtactg	gtgagta	aaccaagtca	ttctgagaat	6960	
agtgtatgcg	gcgaccgagt	tgctttgcc	cggcgtcaat	acgggataat	accgcgccac	7020	
atagcagaac	tttaaaagt	ctcatcattt	gaaaacgtt	ttcggggcga	aaactctcaa	7080	
ggatcttacc	gctgtt	gaga	tccagttcg	tgt	taacccac	tcgtgcaccc	7140
cagcatctt	tactttcacc	agcg	tttctg	gg	tgagcaaa	aacaggaagg	7200
caaaaaagg	aataaggc	gc	acacggaaat	gtt	gaataact	catactcttc	7260
attattgaag	catttatcag	ggttatt	gtc	tcat	gagcgg	atacatattt	7320
agaaaaataa	acaaatagg	gg	gttccgc	gc	catttccc	aaaagtgc	7380
aagaaaccat	tattatcatg	acattaac	ct	ataaaaatag	gcgtatc	acg aggcc	7440
gtc							7443

<210> 22
 <211> 5718
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Sequence of MB1510 genomic integration region

<400>	22					
gagcgccgtt	tggctgaatg	atacaacagt	ctcacttcct	tactgcgtct	ggttgcaaaa	60
acgaagaagc	aaggattccc	ctcgcttctc	atttgccta	tttattatac	acttttttaa	120

gcacatctt	ggcgcttgtt	tcactagact	tgatgcctct	gaatcttgtc	caagtgtcac	180				
ggccgcac	atagacttgt	ccattttca	ccgcttgag	attttccag	agcgggttcg	240				
tttccactc	atctacaatg	gtttgcctt	cgtggctga	gatgaacaaa	atatcaggat	300				
cgatttgct	caattgctca	aggctgacct	cttgataggc	gttatctgac	ttcacagcgt	360				
gtgtaaagcc	tagcattta	aagatttctc	cgtcatagga	tgatgatgta	tgaagctgga	420				
aggaatccgc	tcttgcaacg	ccgagaacga	tgtgcgggtt	ttcatcttcc	ggaagttcgg	480				
cttttagatc	gttgatgact	ttttgtgct	cggcaagctt	ttctttccct	tcatcttctt	540				
tatthaatgc	tttagcaatg	gtcgtaaagc	tgtcgatcgt	ttcgtcatat	gtcgcttcac	600				
ggcttttaa	ttcaatcg	ccccgattt	ttttcagctg	tttataatg	tttttatggc	660				
gctcagcg	agcgatgatt	aaatcaggct	tcaaggaact	gatgacctca	agattgggtt	720				
cgcgcgtgt	gcctacagat	gtgtaatcaa	tggagctgcc	gacaagctt	ttaatcatat	780				
ctttttgtt	gtcatctgcg	atgcccaccc	gcgtaatgcc	gagattgtga	acggcatcca	840				
agaatgaaag	ctcaaggcaca	accacccgct	taggtgtgcc	gcttactgtc	gttttcctt	900				
cttcgtcatg	gatcaactctg	gaatccttag	actcgcttt	gccgcttccg	ttgttattct	960				
ggcttgatg	acagccggat	acaatgaggc	aggcgagcaa	taaaacactc	atgatggcaa	1020				
tcaacttgtt	agaataggtg	cgcgcgtcat	tcttcctttt	ttcagattt	gtaatgagaa	1080				
tcattatcac	atgtaaacact	ataatagcat	ggcttatcat	gtcaatattt	tttttagtaaa	1140				
gaaagctg	tttttactgc	tttctcatga	aagcatcatc	agacacaaat	aagtggatg	1200				
cagcg	tttacc	gtgtcttcga	gacaaaaacg	catggcg	tttagag	1260				
tatcagc	agt	gacataagga	aggagagtgc	tgagataacc	ggacaatttc	1320				
catctgttag	tgcaaattca	atgtcgccga	tattcatgat	aatcgagaaa	acaaagtgcg	1380				
tatcgatatg	aaaatgtcc	tcggcaaaaa	ccgcaagctc	gtgaattcct	ggtgaacatc	1440				
cggcacg	cttacc	atggaaaatc	tgttgacta	aatca	ctc aagca	1500				
gttctgg	taa	aaagtattgc	attagacata	cctcctg	gtacggataa	1560				
tcatgg	tcgt	gtgctccgt	cagcg	cttc	tgat	tttct	1620			
cccgtt	ccta	tcactttacc	atggacggaa	aacaatagc	tactaccatt	cctcctgtt	1680			
ttctctt	caa	tgttctggaa	tctgtttcag	gtacagacga	tcgggtatga	aagaaatata	1740			
gaaaacatg	aa	aggaggaata	tcgacatgaa	accagttgt	aaagagtata	caa	atgacga	1800		
acagctcat	g	aaagatgt	tag	aggaatt	tgca	aaaatgggt	gttgc	aaag	aggatgtata	1860

cgtcttagct cacgacgatg acagaacgga acgcctggct gacaacacga acgccaacac	1920
gatcgaggcc aaagaaacag gtttcaagca cgcggtggaa aatatcttca ataaaaaagg	1980
agacgagctc cgcaataaaa ttcacgaaat cggttttct gaagatgaag ccgctcaatt	2040
tgaaaaacgc ttagatgaag gaaaagtgtct tctctttgtg acagataacg aaaaagtgaa	2100
agcttgggca taaagcaagg aaaaaaccaa aaggccaatg tcggcctttt ggaaaaatgg	2160
cggctttgc ggtgggattt tgcagaatgc cgcaatagga tagcggaaaca ttttcgggtc	2220
tgaatgtccc tcaatttgct attatatttt tgtgataaaat tggaaaaaaat tctcacaaaa	2280
tagaaaaatgg gggtacatag tggatgaaaa aagtgtatgtt agctacggct ttgttttag	2340
gattgactcc agctggcgcg aacgcagctg atttaggccaa ccagacgttggatccaaatg	2400
atggctgggg cgctactcg accggcacga caggcggatc aaaagcatcc tcctcaaattg	2460
tgtataccgt cagcaacaga aaccagcttg tctcggcatt agggaaaggaa acgaacacaa	2520
cgcacaaaaat catttatatc aagggAACGA ttgacatgaa cgtggatgac aatctgaagc	2580
cgttggcct aaatgactat aaagatccgg agtatgattt ggacaaatat ttgaaagcct	2640
atgatcctag cacatggggc aaaaaagagc cgtcgaaac acaagaagaa gcgagagcac	2700
gctctcagaa aaaccaaaaaa gcacgggtca tggtggatat ccctgcaaacc acgacgatcg	2760
tcggttcagg gactaacgct aaagtcgtgg gagggaaactt ccaaatacg agtgataacg	2820
tcattattcg caacattgaa ttccaggatg cctatgacta tttccgcaa tggttgtaaa	2880
acgacggcca gtgaattctg atcaaattggt tcagtggatgg cgaagcgaac acttgatttt	2940
ttaattttct atctttata ggtcattttaga gtatacttata ttgtcctata aactatttag	3000
cagcataata gatttattga ataggtcatt taagttgagc atatttagagg agggaaatct	3060
tggagaaata tttgaagaac ccgagatcta gatcaggatc cgcaacgttc gcagatgctg	3120
ctgaagagat tattaaaaag ctgaaagcaa aaggctatca attggtaact gtatctcagc	3180
ttgaagaagt gaagaagcag agaggctatt gaataaatga gtagaaagcg ccatacgcc	3240
gctttcttt tggaagaaaa tatagggaaa atggtaacttg ttaaaaattc ggaatattta	3300
tacaatatca tatgtatcac attgaaagga gggcctgtct gtccagactg tccgctgtgt	3360
aaaaataagg aataaagggg ggttgacatt attttactga tatgtataat ataatttgta	3420
taagaaaaatg gagggggccct cgaaacgtaa gatgaaacct tagataaaaag tgctttttt	3480
gttgcaattt aagaatttattt aatgttaagc ttaattaaag ataatatctt tgaattgtaa	3540

cgc	ccccctcaa aagtaaaac tacaaaaaaa gaatacgta tatagaataa tgtttgaacc	3600
t	tcttcagat tacaatata ttcggacgga ctctacctca aatgcttac taactataga	3660
a	tgacataca agcacaacct tgaaaattt gaaatataac taccaatgaa cttgttcatg	3720
t	tgaattatcg ctgtatcaa tttctcaat tcaatataa atatgccaat acattgttac	3780
a	aagttagaaat taagacaccc ttgatagcct tactatacct aacatgatgt agtattaaat	3840
g	aatatgtaa atatatttat gataagaagc gacttattt taatcattac atattttct	3900
a	attggaatga ttaagattcc aatagaatag tgtataaatt atttatctt aaaggaggga	3960
c	tgcctaaaaa cgaagaacat taaaaacata tattgcacc gtctaattgaa tttatgaaaa	4020
c	atcatttat cagtttggaa attatgtatt atggagctct gaaaaaaagg agaggataaa	4080
t	gagaaaagg gatcgaaaaa caagtatata ggaggagacc tatttatggc ttcagaaaaa	4140
a	gacgcaggaa aacagtcagc agtaaagctt gttccattgc ttattactgt cgctgtggga	4200
a	ctaattcatct ggttattcc cgctccgtcc ggacttgaac ctaaagctt gcatttgttt	4260
t	gcgatttttgc tcgcaacaat tatcggttt atctccaagc cttgc当地 gggtgcaatt	4320
t	gcaatttttgc cattggcggt tactgcacta actgaaacac tatcaattga ggatacatta	4380
t	agcggattcg ggaataagac cattggctt atcggtatcg catttttgc ttcccgggaa	4440
t	tttatcaaaa ccggtctcg tgcgagaatt tcgtatgtat tcgttgc当地 attcgaaaaa	4500
t	aaaacccttgc gactttctt ttcactgcta ttcaatgtt taatacttgc acctgcttatt	4560
t	ccaagtaata cggcgcgtgc aggaggcatt atatttctt ttatcagatc attatccgaa	4620
t	acattcggat caagccggc aaatgaaaca gagagaaaaa tcggc当地 cttattaaaa	4680
t	accgggttttgc agggaaatct gatcacatct gctatgttcc tgacagcgat ggcggcgaac	4740
t	ccgctgatttgc ccaagctggc ccatgatgtc gcagggtgg acttaacatg gacaagctgg	4800
t	gcaattgc当地 cgattgtacc gggacttgc当地 agcttaatca tcacgc当地 tgtgatttgc	4860
t	aaactgtatc cgccggaaat caaagaaaca ccggatgc当地 cgaaaatcgc aacagaaaaa	4920
t	ctgaaagaaa tgggaccgtt caaaaaatcg gagcttcca tggttatcg ttttttttgc	4980
t	gtgcttgc当地 tgtggatttt tggcggc当地 ttcaacatcg acgctaccac aaccgcatttgc	5040
t	atcggttgc当地 ccgttctt attatcacaat gttctgactt gggatgatca aagaaagaaa	5100
t	cagggc当地 gggatcgc当地 cacttggttt gcggc当地 ttcatgctc当地 caacttcttgc当地	5160
t	aatgaatttgc当地 gcatggtgc当地 ttggttc当地 aatgccc当地 aatcatccgt atcagggttgc当地	5220
t	tcttggatttgc当地 tggcattcat catttttgc当地 gttgtgttatttgc当地 attactctca cttttcttgc当地	5280

gcaagtgcga cagcccacat cagtgcgatg tattcagcat ttttggctgt cgtcgtggca	5340
gcggggcgcac cgccgcgttt agcagcgctg agcctcgctg tcatcagcaa cctgttcggg	5400
tcaacgactc actacggttc tggagcggct ccggctttct tcggagcagg ctacatcccc	5460
caaggcaaat ggtggtccat cggatttatac ctgtcgattt ttcatatcat cgtatggctt	5520
gtgatcggcg gattatggtg gaaagtacta ggaatatggt agaaagaaaa aggccagacgc	5580
ggctgcctt tttttatccc cactccttcg taagaaaatg gattttggaaa aatgagaaaa	5640
ttccctgtga aaaatggtat gatcttaggta gaaaggacgg ctggtgctgt ggtgaaaaag	5700
cgttccatt ttcccttg	5718

<210> 23	
<211> 27	
<212> DNA	
<213> Artificial sequence	
<220>	
<223> Primer 1605	
<400> 23	
gacggccagt gaattcgata aaagtgc	27

<210> 24
<211> 42
<212> DNA
<213> Artificial sequence
<220>
<223> Primer 1606

<220>	
<221> misc_feature	
<222> (13)..(13)	
<223> n is a, c, g, or t	
<220>	
<221> misc_feature	
<222> (16)..(16)	
<223> n is a, c, g, or t	
<400> 24	
ccagatctct atnktnktgt acggagtcta actccccaaag ag	42

<210> 25
<211> 1112
<212> DNA
<213> Nocardiopsis dassonvillei DSM 43235

<400> 25	
gcttttagtt catcgatcgc atcggctgct ccggcccccgt tccccagac ccccgtcgccc	60
gacgacagcg ccgccagcat gaccgaggcg ctcaagcgcg acctcgacct cacctcgccc	120
gaggccgagg agcttctctc ggcgcaggaa gccgcctatcg agaccgacgc cgaggccacc	180
gaggccgagg gcgaggccta cggcggtca ctgttcgaca ccgagaccct cgaactcacc	240
gtgctggtca ccgacgcctc cggcggtca ggcgtcgagg ccaccggagc ccaggccacc	300
gtcgtctccc acggcaccga gggctgacc gaggtcggtgg aggacctcaa cggcgccgag	360
gttcccggaga gcgtcctcggt ctggtacccg gacgtggaga gcgacaccgt cgtggtcgag	420
gtgctggagg gctccgacgc cgacgtcgcc gcctcgctcg ccgacgcccgg tgtggactcc	480
tcctcggtcc gggtgagga ggccgaggag gccccgcagg tctacgcccga catcatcgcc	540
ggcctggcct actacatggg cggccgctgc tccgtcggt tcggccgcgac caacagcgcc	600
ggtcagcccg gtttcgtcac cggcgccac tgcggcaccg tcggcaccgg cgtgaccatc	660
ggcaacggca cggcacctt ccagaactcg gtcttccccg gcaacgacgc cgccttcgtc	720
cggcgaccc ccaacttcac cctgaccaac ctggtctcgcc gctacaactc cggcggtac	780
cagtcggta ccggtaccag ccaggcccg gccggctcggt ccgtgtgccc ctccggctcc	840
accaccggct ggcactgcgg caccatccag gcccccaacc agaccgtgcg ctacccgcag	900
ggcaccgtct actcgctcac cgcaccaac gtgtgcgcgg agccggcga ctccggcggt	960
tcgttcatct ccggctcgca ggcccagggc gtcaccccg gcggtccgg caactgctcc	1020
gtcggcggca cgacctacta ccaggagggtc accccgatga tcaactcctg gggtgtcagg	1080
atccggacct aatcgcatgt tcaatccgct cc	1112

<210> 26
 <211> 48
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Primer 1423

<400> 26	
gcttttagtt catcgatcgc atcggctgct ccggcccccgt tccccag	48

<210> 27
 <211> 45
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Primer 1475

 <400> 27
 ggagcggatt gaacatgcga ttaggtccgg atcctgacac cccag 45

 <210> 28
 <211> 354
 <212> PRT
 <213> Nocardiopsis dassonvillei DSM 43235

 <220>
 <221> PROPEP
 <222> (1)..(166)

 <220>
 <221> mat_peptide
 <222> (167)..(354)

 <400> 28

 Ala Pro Ala Pro Val Pro Gln Thr Pro Val Ala Asp Asp Ser Ala
 -165 -160 -155

 Ala Ser Met Thr Glu Ala Leu Lys Arg Asp Leu Asp Leu Thr Ser
 -150 -145 -140

 Ala Glu Ala Glu Glu Leu Leu Ser Ala Gln Glu Ala Ala Ile Glu
 -135 -130 -125

 Thr Asp Ala Glu Ala Thr Glu Ala Ala Gly Glu Ala Tyr Gly Gly
 -120 -115 -110

 Ser Leu Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp
 -105 -100 -95

 Ala Ser Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gln Ala Thr Val
 -90 -85 -80 -75

 Val Ser His Gly Thr Glu Gly Leu Thr Glu Val Val Glu Asp Leu Asn
 -70 -65 -60

 Gly Ala Glu Val Pro Glu Ser Val Leu Gly Trp Tyr Pro Asp Val Glu
 -55 -50 -45

 Ser Asp Thr Val Val Glu Val Leu Glu Gly Ser Asp Ala Asp Val

-40 -35 -30

Ala Ala Leu Leu Ala Asp Ala Gly Val Asp Ser Ser Ser Val Arg Val
-25 -20 -15

Glu Glu Ala Glu Glu Ala Pro Gln Val Tyr Ala Asp Ile Ile Gly Gly
-10 -5 -1 1 5

Leu Ala Tyr Tyr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr
10 15 20

Asn Ser Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr
25 30 35

Val Gly Thr Gly Val Thr Ile Gly Asn Gly Thr Gly Thr Phe Gln Asn
40 45 50

Ser Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn
55 60 65 70

Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Gln
75 80 85

Ser Val Thr Gly Thr Ser Gln Ala Pro Ala Gly Ser Ala Val Cys Arg
90 95 100

Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn
105 110 115

Gln Thr Val Arg Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr
120 125 130

Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser Gly
135 140 145 150

Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Val
155 160 165

Gly Gly Thr Thr Tyr Tyr Gln Glu Val Thr Pro Met Ile Asn Ser Trp
170 175 180

Gly Val Arg Ile Arg Thr
185

<210> 29
 <211> 498
 <212> DNA
 <213> Nocardiopsis dassonvillei DSM 43235

<400> 29
 gctccggccc ccgtccccca gaccccgta gccgacgaca gcgcgcag catgaccgag 60
 gcgcgtcaagc gcgcacctcg cctcacctcg gccgaggccg aggagcttct ctgcgcgcag 120
 gaagccgcca tcgagaccga cgccgaggcc accgaggccg cggcgaggc ctacggcggc 180
 tcactgttcg acaccgagac cctcgaaactc accgtgctgg tcaccgacgc ctccgcgtc 240
 gaggcggtcg aggccaccgg agcccaggcc accgtcgctc cccacggcac cgagggcctg 300
 accgaggtcg tggaggacct caacggcgcc gaggttcccg agagcgtcct cggtggtag 360
 ccggacgtgg agagcgacac cgtcggtgc gaggtgctgg agggctccga cgccgacgtc 420
 gccgcccgtgc tcgcccacgc cgggtgtggac tcctcctcg tccgggtgga ggaggccgag 480
 gagggccccgc aggtctac 498

<210> 30
 <211> 166
 <212> PRT
 <213> Nocardiopsis dassonvillei DSM 43235

<400> 30

Ala	Pro	Ala	Pro	Val	Pro	Gln	Thr	Pro	Val	Ala	Asp	Asp	Ser	Ala	Ala
1				5				10					15		

Ser	Met	Thr	Glu	Ala	Leu	Lys	Arg	Asp	Leu	Asp	Leu	Thr	Ser	Ala	Glu
				20				25				30			

Ala	Glu	Glu	Leu	Leu	Ser	Ala	Gln	Glu	Ala	Ala	Ile	Glu	Thr	Asp	Ala
					35			40				45			

Glu	Ala	Thr	Glu	Ala	Ala	Gly	Glu	Ala	Tyr	Gly	Gly	Ser	Leu	Phe	Asp
					50			55			60				

Thr	Glu	Thr	Leu	Glu	Leu	Thr	Val	Leu	Val	Thr	Asp	Ala	Ser	Ala	Val
					65			70		75		80			

Glu	Ala	Val	Glu	Ala	Thr	Gly	Ala	Gln	Ala	Thr	Val	Val	Ser	His	Gly
					85				90				95		

Thr Glu Gly Leu Thr Glu Val Val Glu Asp Leu Asn Gly Ala Glu Val
100 105 110

Pro Glu Ser Val Leu Gly Trp Tyr Pro Asp Val Glu Ser Asp Thr Val
115 120 125

Val Val Glu Val Leu Glu Gly Ser Asp Ala Asp Val Ala Ala Leu Leu
130 135 140

Ala Asp Ala Gly Val Asp Ser Ser Ser Val Arg Val Glu Glu Ala Glu
145 150 155 160

Glu Ala Pro Gln Val Tyr
165

<210> 31

<211> 1146

<212> DNA

<213> Artificial sequence

<220>

<223> The DNA sequence coding for the pro-region of SEQ ID NO: 29 fused
in frame to A1918L2 protease tail-variant encoding gene; whole
construct: 10R(proA1918L2).

<400> 31

atgaagaaac	cgttgggaa	aattgtcgca	agcacccgcac	tactcatttc	tgttgcttt	60
agttcatcga	tcgcatcgcc	tgcgccggcc	cccgcccccc	agaccccggt	cggccacgac	120
agcgccgcca	gcatgaccga	ggcgctcaag	cgcgacctcg	acctcaccc	ggccgaggcc	180
gaggagcttc	tctcgccgca	ggaagccgccc	atcgagaccg	acgcccaggc	caccgaggcc	240
gcggcgagg	cctacggcg	ctcactgttc	gacaccgaga	ccctcgaact	caccgtgctg	300
gtcaccgacg	cctccgcgt	cgaggcggtc	gaggccaccc	gagcccaggc	caccgtcg	360
tcccacggca	ccgagggcct	gaccgaggc	gtggaggacc	tcaacggcgc	cgaggttccc	420
gagagcgtcc	tggctggta	ccggacgtg	gagagcgaca	ccgtcgttgt	cgaggtgctg	480
gagggctccg	acgcccacgt	cgcgcctgt	ctcgccgacg	ccggtgtgga	ctccctcctcg	540
gtccgggtgg	aggaggccga	ggaggccccg	caggtctatg	ccgatatcat	tggaggccta	600
gcgtacacaa	tgggtggtcg	ctgcagcgta	ggattgcag	ccacaaatgc	agctggacaa	660
cctggcttcg	tgacagctgg	acattgcggc	cgcgtcggt	cacaggttac	tatcggaat	720
ggaagaggtg	tcttgagca	aagcgtattt	ccggaaatg	atgctgcctt	cgttagaggt	780

acgtccaa	ttacgcttac	taacttagta	tctagataca	acactggcgg	atatgcaact	840	
gtagcagg	tc acaatcaagc	acctattggc	tctagcgtct	gccgctcagg	gtcgactaca	900	
ggatggcatt	gtggaaccat	tcaagctaga	ggtcagagcg	tgagctatcc	tgaaggtacc	960	
gtaacgaaca	tgactcgtac	gactgtatgt	gcagaaccag	gtgactctgg	aggttcatat	1020	
atcagcggta	cgcaagcgca	aggcg	ttaccat	tcaggtggat	ccggtaactg	1080	
ggcacaacgt	tctaccagga	agt	gacaccg	atggtaact	cttggggagt	tagactccgt	1140
acataa						1146	

<210> 32
 <211> 1068
 <212> DNA
 <213> Nocardiopsis Alba DSM 15647

<400> 32	32						
gcgaccggcc	ccctccccca	gtccccacc	ccggatgaag	ccgaggccac	caccatggc	60	
gaggccctcc	agcgcgac	cttcc	ccgcgttcc	ccctctcagg	ccgacgagct	cctcgaggcg	120
caggccgagt	ccttcgagat	cgacgaggcc	gccaccgcgg	ccgcagccga	ctcctacggc	180	
ggctccatct	tcgacaccga	cagcctcacc	ctgaccgtcc	tggtcaccga	cgccctccg	240	
gtcgaggcgg	tcgaggccgc	cggcgcgc	gccaagg	tctgcacgg	catggaggc	300	
ctggaggaga	tcgtcgccga	cctgaac	cg	gccgacgctc	agcccggcgt	cgtggctgg	360
taccccgaca	tccactccga	cacgg	tcgtc	ctcgagg	tcgagg	cggtgcgc	420
gtggactccc	tgctcgccga	cgccgg	gtg	gacaccgc	acgtcaagg	ggagagcacc	480
accgagc	ccgagctgta	cgccgac	atcg	ggtc	atcg	catgggtgg	540
cgctgctcg	tcggcttcgc	ggcc	accaac	gcctccgg	agcccgg	cg	600
ggccactg	gcaccgtcg	cacccgg	tcg	acatcg	acggccagg	cg	660
cg	ttccgtct	tcc	cgccaa	cgactccg	ttcg	ccgc	720
accaac	ctgg	ccgt	caacacc	cggt	ccac	cttcacc	780
gccc	cgatcg	gtc	ccgt	ccat	ccgg	ctcg	840
gtccagg	gatcg	ccgt	ccat	ccgg	ccgt	ccgg	900
accaac	gtcg	ccgt	ccat	ccgg	ccgt	ccgg	960
cagg	cgatcg	ccgt	ccat	ccgg	ccgt	ccgg	1020
gagg	tcacc	ccgt	ccat	ccgg	ccgt	ccgg	1068

<210> 33
<211> 355
<212> PRT
<213> Nocardiopsis Alba DSM 15647

<220>
<221> PROPEP
<222> (1)..(167)

<220>
<221> mat_peptide
<222> (168)..(355)

<400> 33

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Asp Glu Ala Glu
-165 -160 -155

Ala Thr Thr Met Val Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser
-150 -145 -140

Pro Ser Gln Ala Asp Glu Leu Leu Glu Ala Gln Ala Glu Ser Phe
-135 -130 -125

Glu Ile Asp Glu Ala Ala Thr Ala Ala Ala Ala Asp Ser Tyr Gly
-120 -115 -110

Gly Ser Ile Phe Asp Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr
-105 -100 -95

Asp Ala Ser Ala Val Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys
-90 -85 -80

Val Val Ser His Gly Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu
-75 -70 -65 -60

Asn Ala Ala Asp Ala Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile
-55 -50 -45

His Ser Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp
-40 -35 -30

Val Asp Ser Leu Leu Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys
-25 -20 -15

Val	Glu	Ser	Thr	Thr	Glu	Gln	Pro	Glu	Leu	Tyr	Ala	Asp	Ile	Ile	Gly
-10					-5					-1	1				5
Gly	Leu	Ala	Tyr	Thr	Met	Gly	Gly	Arg	Cys	Ser	Val	Gly	Phe	Ala	Ala
	10						15						20		
Thr	Asn	Ala	Ser	Gly	Gln	Pro	Gly	Phe	Val	Thr	Ala	Gly	His	Cys	Gly
	25					30						35			
Thr	Val	Gly	Thr	Pro	Val	Ser	Ile	Gly	Asn	Gly	Gln	Gly	Val	Phe	Glu
	40					45					50				
Arg	Ser	Val	Phe	Pro	Gly	Asn	Asp	Ser	Ala	Phe	Val	Arg	Gly	Thr	Ser
	55					60				65					
Asn	Phe	Thr	Leu	Thr	Asn	Leu	Val	Ser	Arg	Tyr	Asn	Thr	Gly	Gly	Tyr
	70					75			80			85			
Ala	Thr	Val	Ser	Gly	Ser	Ser	Gln	Ala	Ala	Ile	Gly	Ser	Gln	Ile	Cys
		90					95					100			
Arg	Ser	Gly	Ser	Thr	Thr	Gly	Trp	His	Cys	Gly	Thr	Val	Gln	Ala	Arg
	105						110					115			
Gly	Gln	Thr	Val	Ser	Tyr	Pro	Gln	Gly	Thr	Val	Gln	Asn	Leu	Thr	Arg
	120					125				130					
Thr	Asn	Val	Cys	Ala	Glu	Pro	Gly	Asp	Ser	Gly	Gly	Ser	Phe	Ile	Ser
	135					140				145					
Gly	Ser	Gln	Ala	Gln	Gly	Val	Thr	Ser	Gly	Gly	Ser	Gly	Asn	Cys	Ser
	150				155				160			165			
Phe	Gly	Gly	Thr	Thr	Tyr	Tyr	Gln	Glu	Val	Asn	Pro	Met	Leu	Ser	Ser
	170						175				180				
Trp	Gly	Leu	Thr	Leu	Arg	Thr									
		185													

<210> 34
 <211> 43
 <212> DNA
 <213> Artificial sequence

```

<220>
<223> Primer 1421

<400> 34
gttcatcgat cgcatcggt gcgaccggcc ccctccccca gtc 43

<210> 35
<211> 31
<212> DNA
<213> Artificial sequence

<220>
<223> Primer 1604

<400> 35
gcggatccta tcaggtgcgc agggtcagac c 31

<210> 36
<211> 1062
<212> DNA
<213> Nocardiopsis prasina DSM 15648

<400> 36
gccaccggac cgctccccca gtcacccacc ccggaggccg acggcgatctc catgcaggag 60
gcgctccagc gcgacctcg cctgaccccg cttgaggccg atgaactgt ggccgcccag 120
gacaccgcct tcgaggtcga cgaggcccgcc gccgcggccg ccggggacgc ctacggcgcc 180
tccgtcttcg acaccgagac cctgaaactg accgtcctgg tcaccgacgc cgcctcggtc 240
gaggctgtgg aggccaccgg cgcggttacc gaactcgat cctacggcat cgagggcctc 300
gacgagatca tccaggatct caacgcccgc gacgcccgtcc ccggcgttgt cggctggtag 360
ccggacgtgg cgggtgacac cgtcgtcctg gaggtcctgg agggttccgg agccgacgtg 420
agcggcctgc tcgcccacgc cggcgtggac gcctcgccg tcgaggtgac cagcagtgcg 480
cagcccgagc tctacgcccga catcatcgcc ggtctggctt acaccatggg cggccgttgt 540
tcggtcggat tcgcccacca acacgcccgc ggtcagcccg gattcgtcac cgcgggtcac 600
tgtggccgcg tgggcaccca ggtgagcatc ggcaacggcc agggcgttcc cgagcagtcc 660
atcttcccg gcaacgacgc cgccttcgtc cgcggcacgt ccaacttcac gctgaccaac 720
ctggtcagcc gctacaacac cggcggttac gccaccgtcg ccggccacaa ccaggcgccc 780
atcggctcct ccgtctgcgc ctccggctcc accaccggct ggcactgcgg caccatccag 840
gcccgccgcg agtcggtgag ctaccccgag ggcaccgtca ccaacatgac ccggaccacc 900
gtgtgcgcgcg agcccgccga ctccggcgcc tcctacatct ccggcaacca ggcccgaggc 960

```

gtcacacctccg gcggctccgg caactgccgc accggcgaaa ccaccttcta ccaggaggc 1020
accatgg tgaactcctg gggcgccgt ctccggacct aa 1062

<210> 37
<211> 353
<212> PRT
<213> Nocardiopsis prasina DSM 15648

<220>
<221> PROPEP
<222> (1)..(165)

<220>
<221> mat_peptide
<222> (166)..(353)

<400> 37

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala
-165 -160 -155

Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro
-150 -145 -140

Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu
-135 -130 -125

Val Asp Glu Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly
-120 -115 -110

Ser Val Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp
-105 -100 -95 -90

Ala Ala Ser Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu
-85 -80 -75

Val Ser Tyr Gly Ile Glu Gly Leu Asp Glu Ile Ile Gln Asp Leu Asn
-70 -65 -60

Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala
-55 -50 -45

Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val
-40 -35 -30

Ser Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val
 -25 -20 -15 -10

Thr Ser Ser Ala Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu
 -5 -1 1 5

Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn
 10 15 20

Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val
 25 30 35

Gly Thr Gln Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu Gln Ser
 40 45 50 55

Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe
 60 65 70

Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr
 75 80 85

Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser
 90 95 100

Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln
 105 110 115

Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr
 120 125 130 135

Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Asn
 140 145 150

Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly
 155 160 165

Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly
 170 175 180

Val Arg Leu Arg Thr
 185

<210> 38		
<211> 43		
<212> DNA		
<213> Artificial sequence		
<220>		
<223> Primer 1346		
<400> 38		
gttcatcgat cgcatggct gccaccggac cgctcccca gtc		43
<210> 39		
<211> 38		
<212> DNA		
<213> Artificial sequence		
<220>		
<223> Primer 1602		
<400> 39		
gcggatccta ttaggtccgg agacggacgc cccaggag		38
<210> 40		
<211> 1062		
<212> DNA		
<213> Nocardiopsis prasina DSM 15649		
<400> 40		
gccaccggac cactcccca gtcacccacc ccggaggccg acgccgtctc catgcaggag		60
gcgctccagc gcgacctcgg cctgaccccg cttgaggccg atgaactgct ggccgcccag		120
gacaccgcct tcgaggctcg cgaggccgcg gccgaggccg ccggtgacgc ctacggcggc		180
tccgtcttcg acaccgagac cctggaactg accgtcctgg tcaccgactc cgccgcggtc		240
gaggcggtgtgg aggccacccgg cgccgggacc gaactggtct cctacggcat cacgggcctc		300
gacgagatcg tcgaggagct caacgcccgc gacgccgttc ccggcgtggc cggctggtac		360
ccggacgtcg cgggtgacac cgtcgtgctg gaggtcctgg agggttccgg cgccgacgtg		420
ggcggcctgc tcgcccacgc cggcgtggac gcctcggcgg tcgaggtgac caccaccgag		480
cagcccgagc tgtacgcccga catcatcgcc ggtctggcct acaccatggg cggccgctgt		540
tcggtcggct tcgcggccac caacgcccgc ggtcaaccccg ggttcgtcac cgccggtcac		600
tgtggccgcg tgggcaccca ggtgaccatc ggcaacggcc gggcgtt cggcgttcc		660
atcttccccgg gcaacgacgc cgccttcgtc cggaaacgt ccaacttac gctgaccaac		720
ctggtcagcc gctacaacac cggcggctac gccaccgtcg ccggtcacaa ccaggcggcc		780
atcggctcct ccgtctgccc ctccggctcc accaccgggtt ggcactgcgg caccatccag		840

gccccggcgc	agtccgtgag	ctaccccgag	ggcacccgtca	ccaacatgac	gcggaccacc	900
gtgtgcgcgc	agcccgccga	ctccggcggc	tcctacatct	ccggcaacca	ggcccaggc	960
gtcacctccg	gcggctccgg	caactgccgc	accggcggga	ccaccttcta	ccaggaggc	1020
accccatgg	tgaactcctg	ggcggtccgt	ctccggacct	aa		1062

<210> 41
 <211> 353
 <212> PRT
 <213> Nocardiopsis prasina DSM 15649

<220>
 <221> PROPEP
 <222> (1)..(165)

<220>
 <221> mat_peptide
 <222> (166)..(353)

<400> 41

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala
 -165 -160 -155

Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro
 -150 -145 -140

Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu
 -135 -130 -125

Val Asp Glu Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly
 -120 -115 -110

Ser Val Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp
 -105 -100 -95 -90

Ser Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu
 -85 -80 -75

Val Ser Tyr Gly Ile Thr Gly Leu Asp Glu Ile Val Glu Glu Leu Asn
 -70 -65 -60

Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala
 -55 -50 -45

Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val
-40 -35 -30

Gly Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val
-25 -20 -15 -10

Thr Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu
-5 -1 1 5

Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn
10 15 20

Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val
25 30 35

Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser
40 45 50 55

Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe
60 65 70

Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr
75 80 85

Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser
90 95 100

Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln
105 110 115

Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr
120 125 130 135

Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Asn
140 145 150

Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly
155 160 165

Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly
170 175 180

Val Arg Leu Arg Thr
185

<210> 42
<211> 43
<212> DNA
<213> Artificial sequence

<220>
<223> Primer 1603

<400> 42
gttcatcgat cgcatggct gccaccggac cactcccca gtc

43

<210> 43
<211> 353
<212> PRT
<213> Nocardiopsis sp. NRRL 18262

<220>
<221> PROPEP
<222> (1)..(165)

<220>
<221> mat_peptide
<222> (166)..(1059)

<400> 43

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala
-165 -160 -155

Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser
-150 -145 -140

Ala Glu Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu
-135 -130 -125

Val Asp Glu Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly
-120 -115 -110

Ser Val Phe Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr Asp
-105 -100 -95 -90

Ala Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu
-85 -80 -75

Val Ser Tyr Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn

-70	-65	-60
Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala		
-55	-50	-45
Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val		
-40	-35	-30
Ser Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val		
-25	-20	-15
-10		
Thr Thr Ser Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu		
-5	-1	1
5		
Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn		
10	15	20
Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val		
25	30	35
Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser		
40	45	50
55		
Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe		
60	65	70
Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr		
75	80	85
Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser		
90	95	100
Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln		
105	110	115
Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr		
120	125	130
135		
Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr		
140	145	150
Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly		
155	160	165

Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly
170 175 180

Val Arg Leu Arg Thr
185

<210> 44
<211> 1164
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic protease encoding gene

<220>
<221> CDS
<222> (1)..(1164)
<223> Full length protease

<220>
<221> sig_peptide
<222> (1)..(81)

<220>
<221> misc_feature
<222> (82)..(1164)
<223> Propeptide

<220>
<221> mat_peptide
<222> (577)..(1164)

<400> 44
atg aaa aaa ccg ctg gga aaa att gtc gca agc aca gca ctt ctt 45
Met Lys Lys Pro Leu Gly Lys Ile Val Ala Ser Thr Ala Leu Leu
-190 -185 -180

att tca gtg gca ttt agc tca tct att gca tca gca gct aca gga 90
Ile Ser Val Ala Phe Ser Ser Ser Ile Ala Ser Ala Ala Thr Gly
-175 -170 -165

gca tta ccg cag tct ccg aca ccg gaa gca gat gca gtc tca atg 135
Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val Ser Met
-160 -155 -150

caa gaa gca ctg caa aga gat ctt gat ctt aca tca gca gaa gca 180
Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu Ala
-145 -140 -135

gaa gaa ctt ctt gct gca caa gat aca gca ttt gaa gtg gat gaa 225
Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
-130 -125 -120

gca gcg gca gaa gca gca gga gat gca tat ggc ggc tca gtt ttt Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe -115 -110 -105	270
gat aca gaa tca ctt gaa ctt aca gtt ctt gtt aca gat gca gca gca Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ala Ala -100 -95 -90	318
gtt gaa gca gtt gaa gca aca gga gca gga aca gta ctt gtt tca tat Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Val Leu Val Ser Tyr -85 -80 -75	366
gga att gat ggc ctt gat gaa att gtt caa gaa ctg aat gca gct gat Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn Ala Ala Asp -70 -65 -60 -55	414
gct gtt ccg ggc gtt gtt ggc tgg tat ccg gat gtt gct gga gat aca Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr -50 -45 -40	462
gtt gtc ctt gaa gtt ctt gaa gga tca ggc gca gat gtt tca ggc ctg Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Ser Gly Leu -35 -30 -25	510
ctg gca gac gca gga gtc gat gca tca gca gtt gaa gtt aca aca tca Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Thr Ser -20 -15 -10	558
gat caa ccg gaa ctt tat gca gat att att ggc ggc ctg gca tat tat Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu Ala Tyr Tyr -5 -1 1 5 10	606
atg ggc ggc aga tgc agc gtt ggc ttt gca gca aca aat gca tca ggc Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn Ala Ser Gly 15 20 25	654
caa ccg ggc ttt gtt aca gca ggc cat tgc ggc aca gtt ggc aca cca Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr Val Gly Thr Pro 30 35 40	702
gtt tca att ggc aat ggc aaa ggc gtt ttt gaa cga agc att ttt ccg Val Ser Ile Gly Asn Gly Lys Gly Val Phe Glu Arg Ser Ile Phe Pro 45 50 55	750
ggc aat gat tca gca ttt gtt aga ggc aca tca aat ttt aca ctt aca Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser Asn Phe Thr Leu Thr 60 65 70	798
aat ctg gtt tca aga tat aat tca ggc ggc tat gca aca gtt gca ggc Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Ala Thr Val Ala Gly 75 80 85 90	846
cat aat caa gca ccg att ggc tca gca gtt tgc aga tca ggc tca aca His Asn Gln Ala Pro Ile Gly Ser Ala Val Cys Arg Ser Gly Ser Thr 95 100 105	894

aca ggc tgg cat tgc ggc aca att caa gca aga aat caa aca gtt agg	942		
Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn Gln Thr Val Arg			
110	115	120	
 tat ccg caa ggc aca gtt tat agt ctg aca aga aca aca gtt tgt gca	990		
Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr Thr Val Cys Ala			
125	130	135	
 gaa ccg ggc gat tca ggc tca tat att agc ggc act caa gca caa	1038		
Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr Gln Ala Gln			
140	145	150	
 ggc gtt aca tca ggc ggc tca ggc aat tgc agt gct ggc ggc aca aca	1086		
Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Ala Gly Gly Thr Thr			
155	160	165	170
 tat tac caa gaa gtt aat ccg atg ctt agt tca tgg ggc ctt aca ctt	1134		
Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser Trp Gly Leu Thr Leu			
175	180	185	
 aga aca caa tcg cat gtt caa tcc gct cca	1164		
Arg Thr Gln Ser His Val Gln Ser Ala Pro			
190	195		
 <210> 45			
<211> 388			
<212> PRT			
<213> Artificial sequence			
 <220>			
<223> Synthetic Construct			
 <400> 45			
 Met Lys Lys Pro Leu Gly Lys Ile Val Ala Ser Thr Ala Leu Leu			
-190	-185	-180	
 Ile Ser Val Ala Phe Ser Ser Ser Ile Ala Ser Ala Ala Thr Gly			
-175	-170	-165	
 Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val Ser Met			
-160	-155	-150	
 Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu Ala			
-145	-140	-135	
 Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu			
-130	-125	-120	
 Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe			
-115	-110	-105	

Asp	Thr	Glu	Ser	Leu	Glu	Leu	Thr	Val	Leu	Val	Thr	Asp	Ala	Ala	Ala
-100															-90
Val	Glu	Ala	Val	Glu	Ala	Thr	Gly	Ala	Gly	Thr	Val	Leu	Val	Ser	Tyr
-85															-75
Gly	Ile	Asp	Gly	Leu	Asp	Glu	Ile	Val	Gln	Glu	Leu	Asn	Ala	Ala	Asp
-70															-60
Ala	Val	Pro	Gly	Val	Val	Gly	Trp	Tyr	Pro	Asp	Val	Ala	Gly	Asp	Thr
															-45
Val	Val	Leu	Glu	Val	Leu	Glu	Ser	Gly	Ala	Asp	Val	Ser	Gly	Leu	
															-30
Leu	Ala	Asp	Ala	Gly	Val	Asp	Ala	Ser	Ala	Val	Glu	Val	Thr	Thr	Ser
															-10
Asp	Gln	Pro	Glu	Leu	Tyr	Ala	Asp	Ile	Ile	Gly	Gly	Leu	Ala	Tyr	Tyr
-5															10
Met	Gly	Gly	Arg	Cys	Ser	Val	Gly	Phe	Ala	Ala	Thr	Asn	Ala	Ser	Gly
															15
Gln	Pro	Gly	Phe	Val	Thr	Ala	Gly	His	Cys	Gly	Thr	Val	Gly	Thr	Pro
															30
Val	Ser	Ile	Gly	Asn	Gly	Lys	Gly	Val	Phe	Glu	Arg	Ser	Ile	Phe	Pro
															45
Gly	Asn	Asp	Ser	Ala	Phe	Val	Arg	Gly	Thr	Ser	Asn	Phe	Thr	Leu	Thr
															60
Asn	Leu	Val	Ser	Arg	Tyr	Asn	Ser	Gly	Gly	Tyr	Ala	Thr	Val	Ala	Gly
															75
His	Asn	Gln	Ala	Pro	Ile	Gly	Ser	Ala	Val	Cys	Arg	Ser	Gly	Ser	Thr
															95
Thr	Gly	Trp	His	Cys	Gly	Thr	Ile	Gln	Ala	Arg	Asn	Gln	Thr	Val	Arg
															110
															115
															120

Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr Thr Val Cys Ala
125 130 135

Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr Gln Ala Gln
140 145 150

Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Ala Gly Gly Thr Thr
155 160 165 170

Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser Trp Gly Leu Thr Leu
175 180 185

Arg Thr Gln Ser His Val Gln Ser Ala Pro
190 195

<210> 46
<211> 165
<212> PRT
<213> Artificial sequence

<220>
<223> Shuffled propeptide O-2.19

<220>
<221> PROPEP
<222> (1)..(165)

<400> 46

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu
20 25 30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
35 40 45

Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe Asp
50 55 60

Thr Glu Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly

85

90

95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
115 120 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
130 135 140

Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Thr Ser Asp
145 150 155 160

Gln Pro Glu Leu Tyr
165

<210> 47

<211> 166

<212> PRT

<213> Artificial sequence

<220>

<223> Shuffled propeptide G-2.73

<220>

<221> PROPEP

<222> (1)..(166)

<400> 47

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Ser Ser Ala Glu
20 25 30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
35 40 45

Ala Ala Ala Gly Ala Ala Gly Asp Ala Tyr Gly Ser Val Phe Asp
50 55 60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
115 120 125

Val Val Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
130 135 140

Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
145 150 155 160

Glu Gln Pro Glu Leu Tyr
165

<210> 48
<211> 166
<212> PRT
<213> Artificial sequence

<220>
<223> Shuffled propeptide G-1.43

<220>
<221> PROPEP
<222> (1)..(166)

<400> 48

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser Ser Ser Gln
20 25 30

Ala Glu Glu Leu Leu Asp Ala Gln Ala Glu Ser Phe Glu Ile Asp Glu
35 40 45

Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe Asp
50 55 60

Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val

65

70

75

80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
115 120 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
130 135 140

Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
145 150 155 160

Glu Gln Pro Glu Leu Tyr
165

<210> 49
<211> 166
<212> PRT
<213> Artificial sequence

<220>
<223> Shuffled propeptide G-2.6

<400> 49

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu
20 25 30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
35 40 45

Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe Asp
50 55 60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ser Ser Ser Val
65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
115 120 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
130 135 140

Ala Gly Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
145 150 155 160

Glu Gln Pro Glu Leu Tyr
165

<210> 50
<211> 165
<212> PRT
<213> Artificial sequence

<220>
<223> Shuffled propeptide G-2.5

<220>
<221> PROPEP
<222> (1)..(165)

<400> 50

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro Leu Glu
20 25 30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
35 40 45

Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Ser Val Phe Asp
50 55 60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
115 120 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
130 135 140

Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Pro Ala Ala
145 150 155 160

Arg Pro Glu Leu Tyr
165

<210> 51
<211> 166
<212> PRT
<213> Artificial sequence

<220>
<223> Shuffled propeptide G-2.3

<220>
<221> PROPEP
<222> (1)..(166)

<400> 51

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Asp Gly Ala Glu Ala
1 5 10 15

Thr Thr Met Val Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro Ala
20 25 30

Glu Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp
35 40 45

Glu Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe
50 55 60

Asp Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ala Ala
65 70 75 80

Val Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His
85 90 95

Gly Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp
100 105 110

Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr
115 120 125

Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Tyr Ser Leu
130 135 140

Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Pro Ala
145 150 155 160

Ala Gln Pro Glu Leu Tyr
165

<210> 52
<211> 166
<212> PRT
<213> Artificial sequence

<220>
<223> Shuffled propeptide G-1.4

<220>
<221> PROPEP
<222> (1)..(166)

<400> 52

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser Ser Ser Gln
20 25 30

Ala Glu Glu Leu Leu Asp Ala Gln Ala Glu Ser Phe Glu Ile Asp Glu
35 40 45

Ala Ala Ala Ala Ala Ala Asp Ser Tyr Gly Gly Ser Ile Phe Asp
50 55 60

Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
115 120 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
130 135 140

Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
145 150 155 160

Glu Gln Pro Glu Leu Tyr
165

<210> 53
<211> 166
<212> PRT
<213> Artificial sequence

<220>
<223> Shuffled propeptide G-1.2

<220>
<221> PROPEP
<222> (1)...(166)

<400> 53

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu
20 25 30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
35 40 45

Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe Asp
50 55 60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ser Ser Ser Val
65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
115 120 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
130 135 140

Ala Gly Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
145 150 155 160

Glu Gln Pro Glu Leu Tyr
165